

The Complexity of Nash Equilibria in Simple Stochastic Multiplayer Games

Michael Ummels¹ and Dominik Wojtczak^{2,3}

¹ RWTH Aachen University, Germany
E-Mail: ummels@logic.rwth-aachen.de

² CWI, Amsterdam, The Netherlands
E-Mail: d.k.wojtczak@cwi.nl

³ University of Edinburgh, UK

ABSTRACT. *We analyse the computational complexity of finding Nash equilibria in simple stochastic multiplayer games. We show that restricting the search space to equilibria whose payoffs fall into a certain interval may lead to undecidability. In particular, we prove that the following problem is undecidable: Given a game \mathcal{G} , does there exist a pure-strategy Nash equilibrium of \mathcal{G} where player 0 wins with probability 1. Moreover, this problem remains undecidable if it is restricted to strategies with (unbounded) finite memory. However, if mixed strategies are allowed, decidability remains an open problem. One way to obtain a provably decidable variant of the problem is to restrict the strategies to be positional or stationary. For the complexity of these two problems, we obtain a common lower bound of NP and upper bounds of NP and PSPACE respectively.*

1 Introduction

We study *stochastic games* [18] played by multiple players on a finite, directed graph. Intuitively, a play of such a game evolves by moving a token along edges of the graph: Each vertex of the graph is either controlled by one of the players, or it is a *stochastic vertex*. Whenever the token arrives at a non-stochastic vertex, the player who controls this vertex must move the token to a successor vertex; when the token arrives at a stochastic vertex, a fixed probability distribution determines the next vertex. The play ends when it reaches a terminal vertex, in which case each player receives a payoff. In the simplest case, which we discuss here, the possible payoffs of a single play are just 0 and 1 (i.e. each player either wins or loses a given play). However, due to the presence of stochastic vertices, a player's *expected payoff* (i.e. her probability of winning) can be an arbitrary probability.

Stochastic games have been successfully applied in the verification and synthesis of reactive systems under the influence of random events. Such a system is usually modelled as a game between the system and its environment, where the environment's objective is the complement of the system's objective:

the environment is considered hostile. Therefore, traditionally, the research in this area has concentrated on two-player games where each play is won by precisely one of the two players, so-called *two-player, zero-sum games*. However, the system may comprise of several components with independent objectives, a situation which is naturally modelled by a multiplayer game.

The most common interpretation of rational behaviour in multiplayer games is captured by the notion of a *Nash equilibrium* [17]. In a Nash equilibrium, no player can improve her payoff by unilaterally switching to a different strategy. Chatterjee & al. [6] showed that any simple stochastic multiplayer game has a Nash equilibrium, and they also gave an algorithm for computing one. We argue that this is not satisfactory. Indeed, it can be shown that their algorithm may compute an equilibrium where all players lose almost surely (i.e. receive expected payoff 0), while there exist other equilibria where all players win almost surely (i.e. receive expected payoff 1).

In applications, one might look for an equilibrium where as many players as possible win almost surely or where it is guaranteed that the expected payoff of the equilibrium falls into a certain interval. Formulated as a decision problem, we want to know, given a k -player game \mathcal{G} with initial vertex v_0 and two thresholds $\bar{x}, \bar{y} \in [0, 1]^k$, whether (\mathcal{G}, v_0) has a Nash equilibrium with expected payoff at least \bar{x} and at most \bar{y} . This problem, which we call NE for short, is a generalisation of Condon's *SSG Problem* [8] asking whether in a two-player, zero-sum game one of the two players, say player 0, has a strategy to win the game with probability at least $\frac{1}{2}$.

The problem NE comes in several variants, depending on the type of strategies one considers: On the one hand, strategies may be *mixed* (allowing randomisation over actions) or *pure* (not allowing such randomisation). On the other hand, one can restrict to strategies that use (unbounded or bounded) finite memory or even to *stationary* ones (strategies that do not use any memory at all). For the SSG Problem, this distinction is not meaningful since in a two-player, zero-sum simple stochastic game both players have an optimal positional (i.e. both pure and stationary) strategy [8]. However, regarding NE this distinction leads to distinct decision problems, which have to be analysed separately.

Our main result is that NE is undecidable if only pure strategies are considered. In fact, even the following, presumably simpler, problem is undecidable: Given a game \mathcal{G} , decide whether there exists a pure Nash equilibrium where player 0 wins almost surely. Moreover, the problem remains undecidable if one restricts to pure strategies that use (unbounded) finite memory. However, for the general case of arbitrary mixed strategies, decidability remains an open problem.

If one restricts to simpler types of strategies like stationary ones, the problem becomes provably decidable. In particular, for positional strategies the problem becomes NP-complete, and for arbitrary stationary strategies the problem is NP-hard but contained in PSPACE. We also relate the complexity of

the latter problem to the complexity of the infamous *Square Root Sum Problem* (SqrtSum) by providing a polynomial-time reduction from SqrtSum to NE with the restriction to stationary strategies. It is a long-standing open problem whether SqrtSum falls into the polynomial hierarchy; hence, showing that NE for stationary strategies lies inside the polynomial hierarchy would imply a breakthrough in complexity theory.

Let us remark that our game model is rather restrictive: Firstly, players receive a payoff only at terminal vertices. In the literature, a plethora of game models with more complicated modes of winning have been discussed. In particular, the model of a *stochastic parity game* [5, 24] has been investigated thoroughly. Secondly, our model is *turn-based* (i.e. for every non-stochastic vertex there is only one player who controls this vertex) as opposed to *concurrent* [12, 11]. The reason that we have chosen to analyse such a restrictive model is that we are focussing on negative results. Indeed, all our lower bounds hold for (multiplayer versions of) the aforementioned models. Moreover, besides Nash equilibria, our negative results apply to several other solution concepts like subgame perfect equilibria [21, 22] and secure equilibria [4].

For games with *rewards* on transitions [15], the situation might be different: While our lower bounds can be applied to games with rewards under the *average reward* or the *total expected reward* criterion, we leave it as an open question whether this remains true in the case of discounted rewards.

RELATED WORK. Determining the complexity of Nash Equilibria has attracted much interest in recent years. In particular, a series of papers culminated in the result that computing a Nash equilibrium of a two-player game in strategic form is complete for the complexity class PPAD [10, 7]. More in the spirit of our work, Conitzer and Sandholm [9] showed that deciding whether there exists a Nash equilibrium in a two-player game in strategic form where player 0 receives payoff at least x and related decision problems are all NP-hard. For infinite games (without stochastic vertices), (a qualitative version of) the problem NE was studied in [23]. In particular, it was shown that the problem is NP-complete for games with parity winning conditions and even in P for games with Büchi winning conditions.

For stochastic games, most results concern the classical SSG problem: Condon showed that the problem is in $\text{NP} \cap \text{co-NP}$ [8], but it is not known to be in P. We are only aware of two results that are closely related to our problem: Firstly, Etessami & al. [13] investigated Markov decision processes with, e.g., multiple reachability objectives. Such a system can be viewed as a stochastic multiplayer game where all non-stochastic vertices are controlled by one single player. Under this interpretation, one of their results states that NE is decidable in polynomial time for such games. Secondly, Chatterjee & al. [6] showed that the problem of deciding whether a (concurrent) stochastic game with reachability objectives has a positional-strategy Nash equilibrium with payoff at least \bar{x} is NP-complete. We sharpen their hardness result by

showing that the problem remains NP-hard when it is restricted to games with only three players (as opposed to an unbounded number of players) where, additionally, payoffs are assigned at terminal vertices only (cf. Theorem 5 and the subsequent remark).

2 Simple stochastic multiplayer games

The model of a (*two-player, zero-sum*) *simple stochastic game*, introduced by Condon [8], easily generalises to the multiplayer case: Formally, we define a *simple stochastic multiplayer game* (SSMG) as a tuple $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, \Delta, (F_i)_{i \in \Pi})$ such that:

- Π is a finite set of *players* (usually $\Pi = \{0, 1, \dots, k-1\}$);
- V is a finite set of *vertices*;
- $V_i \subseteq V$ and $V_i \cap V_j = \emptyset$ for each $i \neq j \in \Pi$;
- $\Delta \subseteq V \times ([0, 1] \cup \{\perp\}) \times V$ is the *transition relation*;
- $F_i \subseteq V$ for each $i \in \Pi$.

We call a vertex $v \in V_i$ *controlled by player i* and a vertex that is not contained in any of the sets V_i a *stochastic vertex*. We require that a transition is labelled by a probability iff it originates in a stochastic vertex: If $(v, p, w) \in \Delta$ then $p \in [0, 1]$ if v is a stochastic vertex and $p = \perp$ if $v \in V_i$ for some $i \in \Pi$. Moreover, for each pair of a stochastic vertex v and an arbitrary vertex w , we require that there exists precisely one $p \in [0, 1]$ such that $(v, p, w) \in \Delta$. For computational purposes, we require additionally that all these probabilities are rational.

For a given vertex $v \in V$, we denote the set of all $w \in V$ such that there exists $p \in (0, 1] \cup \{\perp\}$ with $(v, p, w) \in \Delta$ by $v\Delta$. For technical reasons, we require that $v\Delta \neq \emptyset$ for all $v \in V$. Moreover, for each stochastic vertex v , the outgoing probabilities must sum up to 1: $\sum_{(p,w):(v,p,w) \in \Delta} p = 1$. Finally, we require that each vertex v that lies in one of the sets F_i is a *terminal (sink) vertex*: $v\Delta = \{v\}$. So if F is the set of all terminal vertices, then $F_i \subseteq F$ for each $i \in \Pi$.

A (*mixed*) *strategy of player i in \mathcal{G}* is a mapping $\sigma : V^*V_i \rightarrow \mathcal{D}(V)$ assigning to each possible *history* $xv \in V^*V_i$ of vertices ending in a vertex controlled by player i a (discrete) probability distribution over V such that $\sigma(xv)(w) > 0$ only if $(v, \perp, w) \in \Delta$. Instead of $\sigma(xv)(w)$, we usually write $\sigma(w \mid xv)$. A (*mixed*) *strategy profile of \mathcal{G}* is a tuple $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ where σ_i is a strategy of player i in \mathcal{G} . Given a strategy profile $\bar{\sigma} = (\sigma_j)_{j \in \Pi}$ and a strategy τ of player i , we denote by $(\bar{\sigma}_{-i}, \tau)$ the strategy profile resulting from $\bar{\sigma}$ by replacing σ_i with τ .

A strategy σ of player i is called *pure* if for each $xv \in V^*V_i$ there exists $w \in v\Delta$ with $\sigma(w \mid xv) = 1$. Note that a pure strategy of player i can be identified with a function $\sigma : V^*V_i \rightarrow V$. A strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ is called *pure* if each σ_i is pure.

A strategy σ of player i in \mathcal{G} is called *stationary* if σ depends only on the current vertex: $\sigma(xv) = \sigma(v)$ for all $xv \in V^*V_i$. Hence, a stationary strategy of player i can be identified with a function $\sigma : V_i \rightarrow \mathcal{D}(V)$. A strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ of \mathcal{G} is called *stationary* if each σ_i is stationary.

We call a pure, stationary strategy a *positional strategy* and a strategy profile consisting of positional strategies only a *positional strategy profile*. Clearly, a positional strategy of player i can be identified with a function $\sigma : V_i \rightarrow V$. More generally, a pure strategy σ is called *finite-state* if it can be implemented by a finite automaton with output or, equivalently, if the equivalence relation $\sim \subseteq V^* \times V^*$ defined by $x \sim y$ if $\sigma(xz) = \sigma(yz)$ for all $z \in V^*V_i$ has only finitely many equivalence classes.¹ Finally, a *finite-state strategy profile* is a profile consisting of finite-state strategies only.

It is sometimes convenient to designate an initial vertex $v_0 \in V$ of the game. We call the tuple (\mathcal{G}, v_0) an *initialised SSMG*. A strategy (strategy profile) of (\mathcal{G}, v_0) is just a strategy (strategy profile) of \mathcal{G} . In the following, we will use the abbreviation SSMG also for initialised SSMGs. It should always be clear from the context if the game is initialised or not.

Given an SSMG (\mathcal{G}, v_0) and a strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$, the *conditional probability of $w \in V$ given the history $xv \in V^*V$* is the number $\sigma_i(w \mid xv)$ if $v \in V_i$ and the unique $p \in [0, 1]$ such that $(v, p, w) \in \Delta$ if v is a stochastic vertex. We abuse notation and denote this probability by $\bar{\sigma}(w \mid xv)$. The probabilities $\bar{\sigma}(w \mid xv)$ induce a probability measure on the space V^ω in the following way: The probability of a basic open set $v_1 \dots v_k \cdot V^\omega$ is 0 if $v_1 \neq v_0$ and the product of the probabilities $\bar{\sigma}(v_j \mid v_1 \dots v_{j-1})$ for $j = 2, \dots, k$ otherwise. It is a classical result of measure theory that this extends to a unique probability measure assigning a probability to every Borel subset of V^ω , which we denote by $\text{Pr}_{v_0}^{\bar{\sigma}}$.

For a set $U \subseteq V$, let $\text{Reach}(U) := V^* \cdot U \cdot V^\omega$. We are mainly interested in the probabilities $p_i := \text{Pr}_{v_0}^{\bar{\sigma}}(\text{Reach}(F_i))$ of reaching the sets F_i . We call p_i the *(expected) payoff of $\bar{\sigma}$ for player i* and the vector $(p_i)_{i \in \Pi}$ the *(expected) payoff of $\bar{\sigma}$* . Another way to define these probabilities is via the *Markov chain $\mathcal{G}^{\bar{\sigma}}$* which is defined as follows: The state set of $\mathcal{G}^{\bar{\sigma}}$ is V^+ (the set of all nonempty sequences of vertices), and the probability of going from state xv to state xvw ($x \in V^*$, $v, w \in V$) is equal to $\bar{\sigma}(w \mid xv)$. Then the expected payoff of $\bar{\sigma}$ for player i can be computed as the probability of reaching a state xv with $v \in F_i$ from state v_0 in $\mathcal{G}^{\bar{\sigma}}$.

DRAWING AN SSMG. When drawing an SSMG as a graph, we will use the following conventions: The initial vertex is marked by an incoming edge that has no source vertex. Vertices that are controlled by a player are depicted as circles, where the player who controls a vertex is given by the label next to it. Stochastic vertices are depicted as diamonds, where the transition

¹In general, this definition is applicable to mixed strategies as well, but for this paper we will identify finite-state strategies with pure finite-state strategies.

probabilities are given by the labels on its outgoing edges (the default being $\frac{1}{2}$). Finally, terminal vertices are generally represented by their associated payoff vector. In fact, we allow arbitrary vectors of rational probabilities as payoffs. This does not increase the power of the model since such a payoff vector can easily be realised by an SSMG consisting of stochastic and terminal vertices only.

3 Nash equilibria

To capture rational behaviour of (selfish) players, John Nash [17] introduced the notion of, what is now called, a *Nash equilibrium*. Formally, given a strategy profile $\bar{\sigma}$, a strategy τ of player i is called a *best response to $\bar{\sigma}$* if τ maximises the expected payoff of player i : $\Pr_{v_0}^{(\bar{\sigma}_{-i}, \tau)}(\text{Reach}(F_i)) \leq \Pr_{v_0}^{(\bar{\sigma}_{-i}, \tau')}$ for all strategies τ' of player i . A Nash equilibrium is a strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ such that each σ_i is a best response to $\bar{\sigma}$. Hence, in a Nash equilibrium no player can improve her payoff by (unilaterally) switching to a different strategy.

Previous research on algorithms for finding Nash equilibria in infinite games has focused on computing *some* Nash equilibrium [6]. However, a game may have several Nash equilibria with different payoffs, and one might not be interested in *any* Nash equilibrium but in one whose payoff fulfils certain requirements. For example, one might look for a Nash equilibrium where certain players win almost surely while certain others lose almost surely. This idea leads us to the following decision problem, which we call NE:²

Given an SSMG (\mathcal{G}, v_0) and thresholds $\bar{x}, \bar{y} \in [0, 1]^{\Pi}$, decide whether there exists a Nash equilibrium of (\mathcal{G}, v_0) with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

For computational purposes, we assume that the thresholds \bar{x} and \bar{y} are vectors of rational numbers. A variant of the problem which omits the thresholds just asks about a Nash equilibrium where some distinguished player, say player 0, wins with probability 1:

Given an SSMG (\mathcal{G}, v_0) , decide whether there exists a Nash equilibrium of (\mathcal{G}, v_0) where player 0 wins almost surely.

Clearly, every instance of the threshold-free variant can easily be turned into an instance of NE (by adding the thresholds $\bar{x} = (1, 0, \dots, 0)$ and $\bar{y} = (1, \dots, 1)$). Hence, NE is, a priori, more general than its threshold-free variant.

Our main concern in this paper are variants of NE where we restrict the type of strategies that are allowed in the definition of the problem: Let PureNE, FinNE, StatNE and PosNE be the problems that arise from NE by restricting the desired Nash equilibrium to consist of pure strategies, finite-state strategies, stationary strategies and positional strategies, respectively.

²In the definition of NE, the ordering \leq is applied componentwise.

In the rest of this paper, we are going to prove upper and lower bounds on the complexity of these problems, where all lower bounds hold for the threshold-free variants, too.

Our first observation is that neither stationary nor pure strategies are sufficient to implement any Nash equilibrium, even if we are only interested in whether a player wins or loses almost surely in the Nash equilibrium. Together with a result from Section 5 (namely Proposition 10), this demonstrates that the problems NE, PureNE, FinNE, StatNE, and PosNE are pairwise distinct problems, which have to be analysed separately.

Proposition 1. There exists an SSMG that has a finite-state Nash equilibrium where player 0 wins almost surely but that has no stationary Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the game \mathcal{G} depicted in Figure 1 played by three players 0, 1 and 2 (with payoffs in this order). Obviously, the following finite-state strategy profile is a Nash equilibrium where player 0 wins almost surely: Player 1 plays from vertex v_2 to vertex v_3 at the first visit of v_2 but leaves the game immediately (by playing to the neighbouring terminal vertex) at all subsequent visits to v_2 ; from vertex v_0 player 1 plays to v_1 ; player 2 plays from vertex v_3 to vertex v_4 at the first visit of v_3 but leaves the game immediately at all subsequent visits to v_3 ; from vertex v_1 player 2 plays to v_2 .

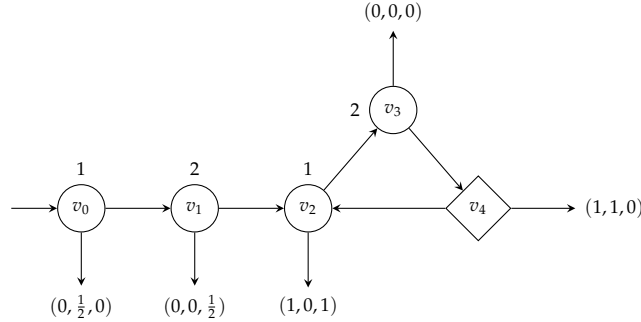


Figure 1. An SSMG with three players

It remains to show that there is no stationary Nash equilibrium of (\mathcal{G}, v_0) where player 0 wins with positive probability. Any stationary Nash equilibrium of (\mathcal{G}, v_0) where player 0 wins with positive probability induces a stationary Nash equilibrium of (\mathcal{G}, v_2) where both players 1 and 2 receive payoff at least $\frac{1}{2}$ since otherwise one of these players could improve her payoff by changing her strategy at v_0 or v_1 . Hence, it suffices to show that (\mathcal{G}, v_2) has no stationary Nash equilibrium where both players 1 and 2 receive payoff at least $\frac{1}{2}$. Assume there exists such an equilibrium and denote by p the probability that player 2 plays from v_3 to v_4 . Since player 1 wins with probability > 0 , it must be the case that $p > 0$. But then, to have a Nash equilibrium, player 1

must play from v_2 to v_3 with probability 1, giving player 2 a payoff of 0, a contradiction. Q.E.D.

Proposition 2. There exists an SSMG that has a stationary Nash equilibrium where player 0 wins almost surely but that has no pure Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the game depicted in Figure 2 played by three players 0, 1 and 2 (with payoffs given in this order). Clearly, the stationary strategy profile where from vertex v_2 player 0 selects both outgoing edges with probability $\frac{1}{2}$ each, player 1 plays from v_0 to v_1 and player 2 plays from v_1 to v_2 is a Nash equilibrium where player 0 wins almost surely. However, for any pure strategy profile where player 0 wins almost surely, either player 1 or player 2 receives payoff 0 and could improve her payoff by switching her strategy at v_0 or v_1 respectively. Q.E.D.

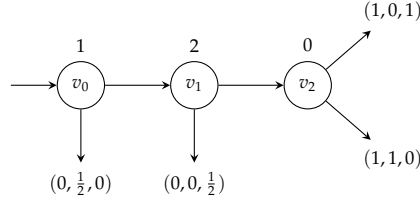


Figure 2. Another SSMG with three players

4 Decidable variants of NE

4.1 Upper bounds

In this section, we show that the problems PosNE and StatNE are contained in the complexity classes NP and PSPACE respectively.

Theorem 3. PosNE is in NP.

Proof. Let (\mathcal{G}, v_0) be an SSMG. Any positional strategy profile of \mathcal{G} can be identified with a mapping $\bar{\sigma} : \bigcup_{i \in \Pi} V_i \rightarrow V$ such that $(v, \perp, \bar{\sigma}(v)) \in \Delta$ for each non-stochastic vertex v , an object whose size is linear in the size of \mathcal{G} . To prove that PosNE is in NP, it suffices to show that we can check in polynomial time whether such a mapping $\bar{\sigma}$ constitutes a Nash equilibrium whose payoff lies in between the given thresholds \bar{x} and \bar{y} .

First, we need to compute the payoff of $\bar{\sigma}$. Let $z_v^i := \Pr_v^{\bar{\sigma}}(\text{Reach}(F_i))$ denote the expected payoff of $\bar{\sigma}$ for player i in (\mathcal{G}, v) , and let $\bar{z}^i = (z_v^i)_{v \in V}$. It is a well-known result of the theory of Markov chains that \bar{z}^i is the optimal solution of the following linear programme:

Minimise $\sum_{v \in V} z_v^i$, subject to:

$$z_v^i \geq 0 \quad \text{for } v \in V,$$

$$\begin{aligned}
z_v^i &= 1 && \text{for } v \in F_i, \\
z_v^i &= \sum_{w \in V} \bar{\sigma}(w \mid v) \cdot z_w^i && \text{for } v \in V \setminus F_i.
\end{aligned}$$

Once we have computed \bar{z}^i , we can check whether $x_i \leq z_{v_0}^i \leq y_i$; this inequality holds for each player $i \in \Pi$ iff the payoff of $\bar{\sigma}$ lies in between \bar{x} and \bar{y} .

To check whether $\bar{\sigma}$ is a Nash equilibrium, we need to compute the numbers $\sup_{\tau} \Pr_{v_0}^{(\bar{\sigma}_{-i}, \tau)}(\text{Reach}(F_i))$ (where τ ranges over every strategy of player i in \mathcal{G}), the maximal payoff that player i can achieve when playing against $\bar{\sigma}_{-i}$. If this payoff is equal to $z_{v_0}^i$, then she cannot gain anything by unilaterally switching to any other strategy. From the theory of *Markov decision process* (cf. [19]), it is well-known that the desired payoff can be computed by the following linear programme over the variables $\bar{r}^i = (r_v^i)_{v \in V}$:

Minimise $\sum_{v \in V} r_v^i$, subject to:

$$\begin{aligned}
r_v^i &\geq 0 && \text{for } v \in V, \\
r_v^i &= 1 && \text{for } v \in F_i, \\
r_v^i &\geq r_w^i && \text{for } v \in V_i \text{ and } w \in v\Delta, \\
r_v^i &= \sum_{w \in V} \bar{\sigma}(w \mid v) \cdot r_w^i && \text{for } v \in V \setminus V_i.
\end{aligned}$$

To check whether $\bar{\sigma}$ is a Nash equilibrium, it suffices to compute for each player i the optimal solution \bar{r}^i and to check whether $r_{v_0}^i = z_{v_0}^i$.

Since linear programmes can be solved in polynomial time and both programmes are of size polynomial in the size of the game, all these checks can be carried out in polynomial time. Q.E.D.

To prove the decidability of StatNE, we appeal to results established for the *Existential Theory of the Reals*, $\text{ExTh}(\mathfrak{R})$, the set of all existential first-order sentences (over the appropriate signature) that hold in $\mathfrak{R} := (\mathbb{R}, +, \cdot, 0, 1, \leq)$. The best known upper bound for the complexity of the associated decision problem is PSPACE [3, 20], which leads to the following theorem.

Theorem 4. StatNE is in PSPACE.

Proof. Instead of giving a deterministic polynomial-space algorithm for StatNE, we give a nondeterministic one. Since $\text{PSPACE} = \text{NPSPACE}$, this implies that StatNE is in PSPACE. On input $\mathcal{G}, v_0, \bar{x}, \bar{y}$, the algorithm starts by guessing a set $S \subseteq V \times V$ and proceeds by computing, for each player i , the set R_i of vertices from where the set F_i is reachable in the graph $G = (V, S)$, a computation which can be carried out in polynomial time. Note that if S is the *support* of a stationary strategy profile $\bar{\sigma}$, i.e. $S = \{(v, w) \in V \times V : \bar{\sigma}(w \mid v) > 0\}$, then R_i is precisely the set of vertices v such that $\Pr_{v_0}^{\bar{\sigma}}(\text{Reach}(F_i)) > 0$. Finally, the algorithm evaluates an existential first-order sentence ψ , which can be computed in polynomial time from $(\mathcal{G}, v_0), \bar{x}, \bar{y}, S$ and $(R_i)_{i \in \Pi}$, over \mathfrak{R} and returns the answer to this query.

It remains to describe a suitable sentence ψ . Let $\bar{\alpha} = (\alpha_{vw})_{v,w \in V}$, $\bar{r} = (r_v^i)_{i \in \Pi, v \in V}$ and $\bar{z} = (z_v^i)_{i \in \Pi, v \in V}$ be three sets of variables, and let $V_* = \bigcup_{i \in \Pi} V_i$ be the set of all non-stochastic vertices. The formula

$$\begin{aligned} \varphi(\bar{\alpha}) := & \bigwedge_{v \in V_*} \left(\bigwedge_{w \in v\Delta} \alpha_{vw} \geq 0 \wedge \bigwedge_{w \in V \setminus v\Delta} \alpha_{vw} = 0 \wedge \sum_{w \in v\Delta} \alpha_{vw} = 1 \right) \wedge \\ & \bigwedge_{\substack{v \in V \setminus V_* \\ w \in V}} \alpha_{vw} = p_{vw} \wedge \bigwedge_{(v,w) \in S} \alpha_{vw} > 0 \wedge \bigwedge_{(v,w) \notin S} \alpha_{vw} = 0, \end{aligned}$$

where p_{vw} is the unique number such that $(v, p_{vw}, w) \in \Delta$, states that the mapping $\bar{\sigma} : V \rightarrow \mathcal{D}(V)$ defined by $\bar{\sigma}(w \mid v) = \alpha_{vw}$ constitutes a valid stationary strategy profile of \mathcal{G} whose support is S . Provided that $\varphi(\bar{\alpha})$ holds in \mathfrak{R} , the formula

$$\eta_i(\bar{\alpha}, \bar{z}) := \bigwedge_{v \in F_i} z_v^i = 1 \wedge \bigwedge_{v \in V \setminus R_i} z_v^i = 0 \wedge \bigwedge_{v \in V \setminus F_i} z_v^i = \sum_{w \in v\Delta} \alpha_{vw} z_w^i$$

states that $z_v^i = \Pr_v^{\bar{\sigma}}(\text{Reach}(F_i))$ for each $v \in V$, where $\bar{\sigma}$ is defined as above. Again, this follows from a well-known results about Markov chains, namely that the vector of the aforementioned probabilities is the unique solution to the given system of equations. Finally, the formula

$$\vartheta_i(\bar{\alpha}, \bar{r}) := \bigwedge_{v \in V} r_v^i \geq 0 \wedge \bigwedge_{v \in F_i} r_v^i = 1 \wedge \bigwedge_{\substack{v \in V_i \\ w \in v\Delta}} r_v^i \geq r_w^i \wedge \bigwedge_{v \in V \setminus V_i} r_v^i = \sum_{w \in v\Delta} \alpha_{vw} r_w^i$$

states that \bar{r} is a solution of the linear programme for computing the maximal payoff that player i can achieve when playing against the strategy profile $\bar{\sigma}_{-i}$. In particular, the formula is fulfilled if $r_v^i = \sup_{\tau} \Pr_v^{(\bar{\sigma}_{-i}, \tau)}(\text{Reach}(F_i))$ (where τ ranges over *every* strategy of player i), and every other solution is greater than this one (in each component).

The desired sentence ψ is the existential closure of the conjunction of φ and, for each player i , the formulae η_i and ϑ_i combined with formulae stating that player i cannot improve her payoff and that the expected payoff for player i lies in between the given thresholds:

$$\psi := \exists \bar{\alpha} \exists \bar{r} \exists \bar{z} \left(\varphi(\bar{\alpha}) \wedge \bigwedge_{i \in \Pi} (\eta_i(\bar{\alpha}, \bar{z}) \wedge \vartheta_i(\bar{\alpha}, \bar{r}) \wedge r_{v_0}^i \leq z_{v_0}^i \wedge x_i \leq z_{v_0}^i \leq y_i) \right)$$

It follows that ψ holds in \mathfrak{R} iff (\mathcal{G}, v_0) has a stationary Nash equilibrium $\bar{\sigma}$ with payoff at least \bar{x} and at most \bar{y} whose support is S . Consequently, the algorithm is correct. Q.E.D.

4.2 Lower bounds

Having shown that PosNE and StatNE are in NP and PSPACE respectively, the natural question arises whether there is a polynomial-time algorithm for PosNE or StatNE. The following theorem shows that this is not the case (unless, of course, $P = NP$) since both problems are NP-hard. Moreover, both problems are already NP-hard for games with only two players.

Theorem 5. PosNE and StatNE are NP-hard, even for games with only two players.

Proof. The proof is by reduction from SAT. Let $\varphi = C_1 \wedge \dots \wedge C_m$ be a formula in conjunctive normal form over propositional variables X_1, \dots, X_n . Our aim is to construct a two-player SSMG $(\mathcal{G}_\varphi, v_0)$ such that the following statements are equivalent:

1. φ is satisfiable;
2. $(\mathcal{G}_\varphi, v_0)$ has a positional Nash equilibrium with payoff $(1, \frac{1}{2})$;
3. $(\mathcal{G}_\varphi, v_0)$ has a stationary Nash equilibrium with payoff $(1, \frac{1}{2})$.

Provided that the game can be constructed in polynomial time, the equivalence of 1. and 2. establishes a polynomial-time reduction from SAT to PosNE, whereas the equivalence of 1. and 3. establishes one from SAT to StatNE. The game \mathcal{G}_φ is depicted in Figure 3 and played by players 0 and 1. The game proceeds from the initial vertex v_0 to X_i or \bar{X}_i with probability $\frac{1}{2^{i+1}}$ each, and there is an edge from vertex C_j to vertex X_i or \bar{X}_i iff X_i or $\neg X_i$ respectively occurs in the clause C_j . Also, from \top -labelled vertices player 1 can “leave the game” by moving to a terminal vertex with payoff $(0, 1)$. Obviously, the game \mathcal{G}_φ can be constructed from φ in polynomial time. It remains to show that 1.–3. are equivalent.

(1. \Rightarrow 2.) Assume that $\alpha : \{X_1, \dots, X_n\} \rightarrow \{\text{true}, \text{false}\}$ is a satisfying assignment of φ . In the positional Nash equilibrium of (\mathcal{G}, v_0) , player 0 moves from a literal L (i.e. $L = X_i$ or $L = \bar{X}_i$ for some $i = 1, \dots, n$) to the \top -labelled vertex iff L is mapped to true by α , and player 1 moves from vertex C_j to a (fixed) literal L that is contained in C_j and mapped to true by α (which is possible since α is a satisfying assignment). At \top -labelled vertices, player 1 never leaves the game. Obviously, player 0 wins almost surely with this strategy profile. For player 1, the payoff is

$$\frac{1}{2^{n+1}} + \sum_{i=1}^n \frac{1}{2^{i+1}} = \frac{1}{2^{n+1}} + \frac{1}{2} \left(\sum_{i=1}^n \frac{1}{2^i} \right) = \frac{1}{2^{n+1}} + \frac{1}{2} \left(1 - \frac{1}{2^n} \right) = \frac{1}{2},$$

where the first summand is the probability of going from the initial vertex to φ , from where player 1 wins almost surely since from every clause vertex she plays to a “true” literal. Obviously, changing her strategy cannot give her a better payoff. Therefore, we have a Nash equilibrium.

(2. \Rightarrow 3.) Obvious.

(3. \Rightarrow 1.) Let $\bar{\sigma} = (\sigma_0, \sigma_1)$ be a stationary Nash equilibrium of $(\mathcal{G}_\varphi, v_0)$ with payoff $(1, \frac{1}{2})$. Our first aim is to show that σ_0 is actually a positional strategy. Towards a contradiction, assume that there exists a literal L such that $\sigma_0(L)$ assigns probability $0 < q < 1$ to the neighbouring \top -labelled vertex. Since player 0 wins almost surely, player 1 never leaves the game. Hence, the expected payoff for player 1 from vertex L (i.e. in the game (\mathcal{G}_φ, L)) is precisely q . However, if she left the game at the \top -labelled vertex, she

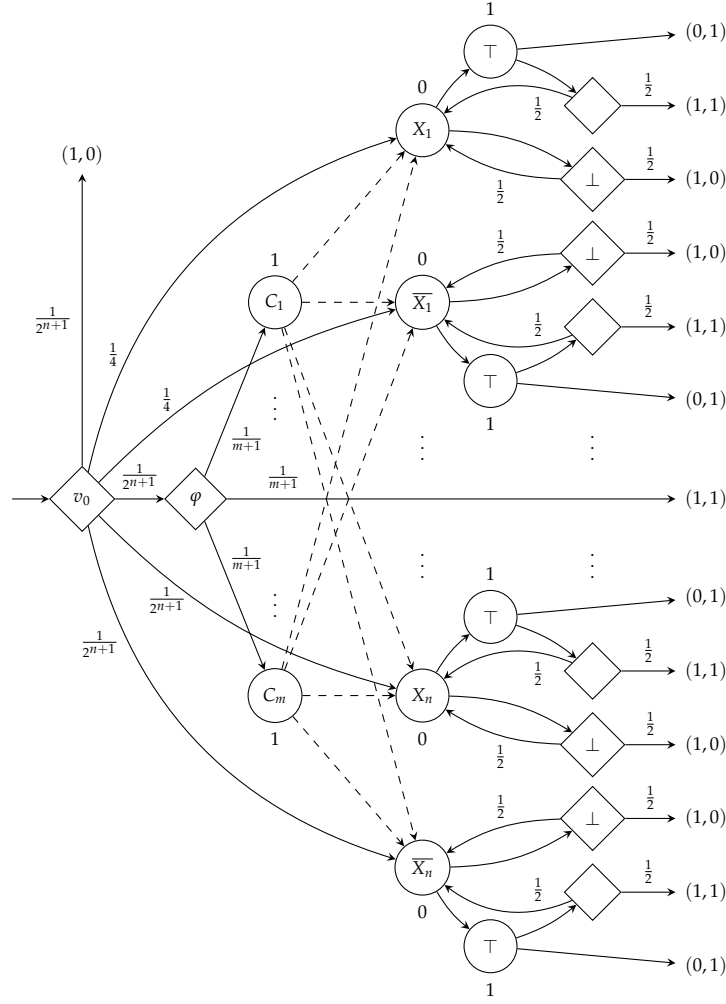


Figure 3. Reducing SAT to PosNE and StatNE.

would receive payoff $\frac{2q}{1+q} > q$. Therefore, $\bar{\sigma}$ is not a Nash equilibrium, a contradiction.

Knowing that σ_0 is a positional strategy, we can define a *pseudo assignment* $\alpha : \{X_1, \neg X_1, \dots, X_n, \neg X_n\} \rightarrow \{\text{true}, \text{false}\}$ by setting $\alpha(L) = \text{true}$ if σ_1 prescribes to go from vertex L to the neighbouring \top -labelled vertex. Our next aim is to show that α is actually an assignment: $\alpha(X_i) = \text{true} \Leftrightarrow \alpha(\neg X_i) = \text{false}$. To see this, note that we can compute player 1's expected payoff as follows:

$$\frac{1}{2} = \frac{p}{2^{n+1}} + \sum_{i=1}^n \frac{a_i}{2^{i+1}}, \quad a_i = \begin{cases} 0 & \text{if } \alpha(X_i) = \alpha(\neg X_i) = \text{false}, \\ 1 & \text{if } \alpha(X_i) \neq \alpha(\neg X_i), \\ 2 & \text{if } \alpha(X_i) = \alpha(\neg X_i) = \text{true}, \end{cases}$$

where p is the expected payoff for player 1 from vertex φ . By the construction

of \mathcal{G}_p , we have $p > 0$, and the equality only holds if $p = 1$ and $a_i = 1$ for all $i = 1, \dots, n$, which proves that α is an assignment.

Finally, we claim that α is a satisfying assignment. If this were not the case, there would exist a clause C such that player 1's expected payoff from vertex C is 0 and therefore $p < 1$, where p is defined as above. This is a contradiction to the fact that $p = 1$, as we have shown above. Q.E.D.

Remark. The reduction in the proof of Theorem 5 can be modified to demonstrate NP-hardness of the threshold-free variants of PosNE and StatNE, albeit at the expense of adding one more player to the game.

It follows from Theorems 3 and 5 that PosNE is NP-complete. For StatNE, we have provided an NP lower bound and a PSPACE upper bound, but the exact complexity of the problem remains unclear. Towards gaining more insight into the problem StatNE, we relate its complexity to the complexity of the *Square Root Sum Problem* (SqrtSum), the problem of deciding, given numbers $d_1, \dots, d_n, k \in \mathbb{N}$, whether $\sum_{i=1}^n \sqrt{d_i} \geq k$. Recently, it was shown that SqrtSum belongs to the 4th level of the *counting hierarchy* [1], which is a slight improvement over the previously known PSPACE upper bound. However, it is an open question since the 1970s whether SqrtSum falls into the polynomial hierarchy [16, 14]. We identify a polynomial-time reduction from SqrtSum to StatNE.³ Hence, StatNE is at least as hard as SqrtSum, and showing that StatNE resides inside the polynomial hierarchy would imply a major breakthrough in understanding the complexity of numerical computation.

Theorem 6. SqrtSum is polynomial-time reducible to StatNE.

Proof. Given an instance (d_1, \dots, d_n, k) of SqrtSum, we construct an SSMG (\mathcal{G}, v_0) played by players 0, 1, 2, 3 (with payoffs given in this order) such that $\sum_{i=1}^n \sqrt{d_i} \geq k$ iff (\mathcal{G}, v_0) has a stationary Nash equilibrium where player 0 wins almost surely.

In order to state our reduction, let us first examine the game $\mathcal{G}(p)$, where $p \in [\frac{1}{2}, 1)$, which is depicted in Figure 4 (b).

Claim 7. The maximal payoff player 3 can receive in a stationary Nash equilibrium of $(\mathcal{G}(p), s)$ is $\frac{\sqrt{2-2p-p+1}}{2p+2}$.

Proof. Let $\bar{\sigma}$ be any stationary strategy profile of $(\mathcal{G}(p), s)$. We denote by x_1 and x_2 the probabilities that player 0 stays inside the gadget at vertex s_1 and vertex s_2 respectively. Consequently, the probabilities of *eventually leaving* the gadget at vertex s_1 and vertex s_2 are given by $p_1(x_1, x_2) := \frac{p(1-x_1)}{1-x_1x_2p^2}$ and $p_2(x_1, x_2) := \frac{p(1-x_2)}{1-x_1x_2p^2}$ respectively. Note that if $x_1 = 0$, then $\bar{\sigma}$ is a Nash equilibrium where player 3 receives payoff $\leq 1 - p \leq \frac{\sqrt{2-2p-p+1}}{2p+2}$. Hence, let us assume that $x_1 > 0$ and look for a Nash equilibrium where player 3 receives

³Some authors define SqrtSum with \leq instead of \geq . With this definition, we would reduce from the complement of SqrtSum instead.

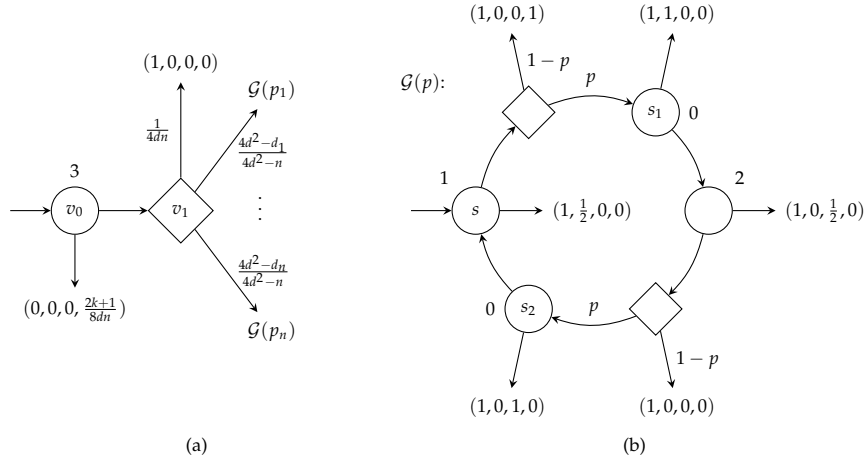


Figure 4. Reducing SqrtSum to StatNE.

payoff $> 1 - p$. For this, it must be the case that $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ since otherwise player 1 or player 2 could improve her payoff by moving out of the gadget, where they would get payoff $\frac{1}{2}$ immediately (and player 3 would receive payoff $\leq 1 - p$). Vice versa, if $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ then $\bar{\sigma}$ is obviously a Nash equilibrium. Hence, to determine the maximum payoff for player 3 in a stationary Nash equilibrium, we have to maximise $\frac{1-p}{1-x_1x_2p^2}$, the expected payoff for player 3, under the constraints $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ and $0 \leq x_1, x_2 \leq 1$. We claim that the maximum is reached only if $x_1 = x_2$; if, for example, $x_1 > x_2$ then we can achieve a higher payoff for player 3 by setting $x'_2 := x_1$, and the constraints are still satisfied:

$$\frac{p(1-x'_2)}{1-x_1x'_2p^2} = \frac{p(1-x_1)}{1-x_1x'_2p^2} = \frac{p(1-x_1)}{1-x_1^2p^2} \geq \frac{p(1-x_1)}{1-x_1x_2p^2} \geq \frac{1}{2}$$

Hence, in fact, we have to maximise $\frac{1-p}{1-x^2p^2}$ under the constraints $\frac{p(1-x)}{1-x^2p^2} \geq \frac{1}{2}$ and $0 \leq x \leq 1$, i.e. under $p^2x^2 - 2px + 2p - 1 \geq 0$ and $0 \leq x \leq 1$. The roots of $p^2x^2 - 2px + 2p - 1$ are $\frac{1 \pm \sqrt{2-2p}}{p}$, but $\frac{1+\sqrt{2-2p}}{p}$ is always greater than 1 for $p \in [0, 1)$. Hence, any solution must be less than $x := \frac{1-\sqrt{2-2p}}{p}$. In fact, we always have $0 \leq x < 1$ for $p \in (\frac{1}{2}, 1)$. Therefore, x is the optimal solution, and the maximal payoff for player 3 is indeed $\frac{1-p}{1-x^2p^2} = \frac{\sqrt{2-2p}-p+1}{2p+2}$.

Finally, we can setup our reduction. Let (d_1, \dots, d_n, k) be an instance of SqrtSum where, without loss of generality, $n > 0$, $d_i > 0$ for each $i = 1, \dots, n$, and $k \leq d := \sum_{i=1}^n d_i$. Define $p_i := 1 - \frac{d_i}{2d^2}$ for $i = 1, \dots, n$. Note that $p_i \in [\frac{1}{2}, 1)$ since $0 < d_i \leq d \leq d^2$. For the reduction, we use n copies of the game $G(p)$, where in the i th copy we set p to p_i . The complete game G is depicted in Figure 4 (a); it can obviously be constructed in polynomial time.

By the above claim, the maximal payoff player 3 can get in a stationary

Nash equilibrium of $(\mathcal{G}(p_i), s)$ is

$$\frac{\sqrt{2-2p_i} - p_i + 1}{2p_i + 2} = \frac{\frac{1}{d}\sqrt{d_i} - (1 - \frac{d_i}{2d^2}) + 1}{4 - \frac{d_i}{d^2}} = \frac{d\sqrt{d_i} + \frac{d_i}{2}}{4d^2 - d_i}.$$

Consequently, the maximal payoff player 3 can get in a stationary Nash equilibrium of (\mathcal{G}, v_1) is

$$\sum_{i=1}^n \frac{4d^2 - d_i}{4d^2n} \cdot \frac{d\sqrt{d_i} + \frac{d_i}{2}}{4d^2 - d_i} = \sum_{i=1}^n \frac{\sqrt{d_i}}{4dn} + \sum_{i=1}^n \frac{d_i}{8d^2n} = \sum_{i=1}^n \frac{\sqrt{d_i}}{4dn} + \frac{1}{8dn}.$$

Let us fix a stationary Nash equilibrium $\bar{\sigma}$ of (\mathcal{G}, v_1) with this payoff for player 3.

Now, if $\sum_{i=1}^n \sqrt{d_i} \geq k$, then also $\sum_{i=1}^n \frac{\sqrt{d_i}}{4dn} + \frac{1}{8dn} \geq \frac{2k+1}{8dn}$, and $\bar{\sigma}$ can be extended to a stationary Nash equilibrium of (\mathcal{G}, v_0) where player 0 wins almost surely by setting $\bar{\sigma}(v_1 | v_0) = 1$. On the other hand, if $\sum_{i=1}^n \sqrt{d_i} < k$, then also $\sum_{i=1}^n \frac{\sqrt{d_i}}{4dn} + \frac{1}{8dn} < \frac{2k+1}{8dn}$, and in every stationary Nash equilibrium of (\mathcal{G}, v_0) player 3 leaves the game at v_0 , which gives payoff 0 to player 0. Q.E.D.

5 Undecidable variants of NE

5.1 Pure-strategy equilibria

In this section, we show that the problem PureNE is undecidable by exhibiting a reduction from an undecidable problem about *two-counter machines*. Our construction is inspired by a construction used by Brázdil & al. [2] to prove the undecidability of stochastic games with branching-time winning conditions.

A two-counter machine \mathcal{M} is given by a list of instructions ι_1, \dots, ι_m where each instruction is one of the following:

- “inc(j); goto k ” (increment counter j by 1 and go to instruction number k);
- “zero(j) ? goto k : dec(j); goto l ” (if the value of counter j is zero, go to instruction number k ; otherwise, decrement counter j by one and go to instruction number l);
- “halt” (stop the computation).

Here j ranges over 1, 2 (the two counters), and $k \neq l$ range over $1, \dots, m$. A configuration of \mathcal{M} is a triple $C = (i, c_1, c_2) \in \{1, \dots, m\} \times \mathbb{N} \times \mathbb{N}$, where i denotes the number of the current instruction and c_j denotes the current value of counter j . A configuration C' is the *successor* of configuration C , denoted by $C \vdash C'$, if it results from C by executing instruction ι_i ; a configuration $C = (i, c_1, c_2)$ with $\iota_i = \text{“halt”}$ has no successor configuration. Finally, the *computation* of \mathcal{M} is the unique maximal sequence $\rho = \rho(0)\rho(1)\dots$ such that $\rho(0) \vdash \rho(1) \vdash \dots$ and $\rho(0) = (1, 0, 0)$ (the *initial configuration*). Note that ρ is either infinite, or it ends in a configuration $C = (i, c_1, c_2)$ such that $\iota_i = \text{“halt”}$.

The *halting problem* is to decide, given a machine \mathcal{M} , whether the computation of \mathcal{M} is finite. It is well-known that two-counter machines are Turing powerful, which makes the halting problem and its dual, the *non-halting problem*, undecidable.

Theorem 8. PureNE is undecidable.

In order to prove Theorem 8, we show that one can compute from a two-counter machine \mathcal{M} an SSMG (\mathcal{G}, v_0) with nine players such that the computation of \mathcal{M} is infinite iff (\mathcal{G}, v_0) has a pure Nash equilibrium where player 0 wins almost surely. This establishes a reduction from the non-halting problem to PureNE.

The game \mathcal{G} is played by player 0 and eight other players A_j^t and B_j^t , indexed by $j \in \{1, 2\}$ and $t \in \{0, 1\}$. Let $\Gamma = \{\text{init}, \text{inc}(j), \text{dec}(j), \text{zero}(j) : j = 1, 2\}$. If \mathcal{M} has instructions ι_1, \dots, ι_m , then for each $i \in \{1, \dots, m\}$, each $\gamma \in \Gamma$, each $j \in \{1, 2\}$ and each $t \in \{0, 1\}$, the game \mathcal{G} contains the gadgets $S_{i,\gamma}^t$, $I_{i,\gamma}^t$ and $C_{j,\gamma}^t$, which are depicted in Figure 5. In the figure, squares represent terminal vertices (the edge leading from a terminal vertex to itself being implicit), and the labelling indicates which players win at the respective vertex. Moreover, the dashed edge inside $C_{j,\gamma}^t$ is present iff $\gamma \notin \{\text{init}, \text{zero}(j)\}$. The initial vertex v_0 of \mathcal{G} is the black vertex inside the gadget $S_{1,\text{init}}^0$.

For any pure strategy profile $\bar{\sigma}$ of \mathcal{G} where player 0 wins almost surely, let $x_0 v_0 \prec x_1 v_1 \prec x_2 v_2 \prec \dots$ ($x_i \in V^*$, $v \in V$, $x_0 = \varepsilon$) be the (unique) sequence of all consecutive histories such that, for each $n \in \mathbb{N}$, v_n is a black vertex and $\Pr_{v_0}^{\bar{\sigma}}(x_n v_n \cdot V^\omega) > 0$. Additionally, let $\gamma_0, \gamma_1, \dots$ be the corresponding sequence of instructions, i.e. $\gamma_n = \gamma$ for the unique instruction γ such that v_n lies in one of the gadgets $S_{i,\gamma}^t$ (where $t = n \bmod 2$). For each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, we define two conditional probabilities a_j^n and p_j^n as follows:

$$a_j^n := \Pr_{v_0}^{\bar{\sigma}}(\text{Reach}(F_{A_j^{n \bmod 2}}) \mid x_n v_n \cdot V^\omega)$$

and

$$p_j^n := \Pr_{v_0}^{\bar{\sigma}}(\text{Reach}(F_{A_j^{n \bmod 2}}) \mid x_n v_n \cdot V^\omega \setminus x_{n+2} v_{n+2} \cdot V^\omega).$$

Finally, for each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, we define an ordinal number $c_j^n \leq \omega$ as follows: After the history $x_n v_n$, with probability $\frac{1}{8}$ the play proceeds to the vertex controlled by player 0 in the counter gadget C_{j,γ_n}^t (where $t = n \bmod 2$). The number c_j^n is defined to be the maximal number of subsequent visits to the grey vertex inside this gadget (where $c_j^n = \omega$ if, on one path, the grey vertex is visited infinitely often). Note that, by the construction of $C_{j,\gamma}^t$, it holds that $c_j^n = 0$ if $\gamma_n = \text{zero}(j)$ or $\gamma_n = \text{init}$.

Lemma 9. Let $\bar{\sigma}$ be a pure strategy profile of (\mathcal{G}, v_0) where player 0 wins

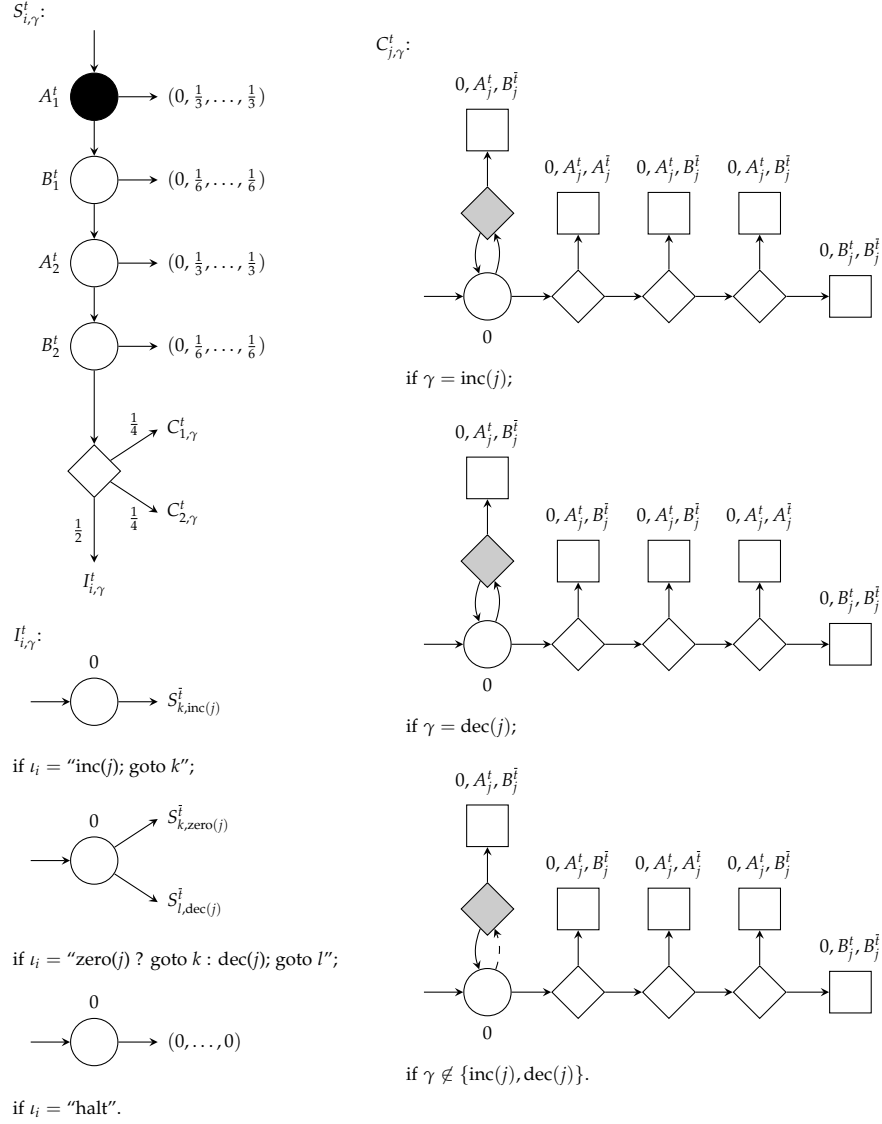


Figure 5. Simulating a two-counter machine.

almost surely. Then $\bar{\sigma}$ is a Nash equilibrium if and only if

$$c_j^{n+1} = \begin{cases} 1 + c_j^n & \text{if } \gamma_{n+1} = \text{inc}(j), \\ c_j^n - 1 & \text{if } \gamma_{n+1} = \text{dec}(j), \\ c_j^n = 0 & \text{if } \gamma_{n+1} = \text{zero}(j), \\ c_j^n & \text{otherwise} \end{cases} \quad (1)$$

for all $j \in \{1, 2\}$ and $n \in \mathbb{N}$.

Here $+$ and $-$ denote the usual addition and subtraction of ordinal numbers respectively (satisfying $1 + \omega = \omega - 1 = \omega$). The proof of Lemma 9 goes through several claims. In the following, let $\bar{\sigma}$ be a pure strategy profile of (\mathcal{G}, v_0) where player 0 wins almost surely. The first claim gives a necessary and sufficient condition on the probabilities a_j^n for $\bar{\sigma}$ to be a Nash equilibrium.

Claim. The profile $\bar{\sigma}$ is a Nash equilibrium iff $a_j^n = \frac{1}{3}$ for all $j \in \{1, 2\}$ and $n \in \mathbb{N}$.

Proof. (\Rightarrow) Assume that $\bar{\sigma}$ is a Nash equilibrium. Clearly, this implies that $a_j^n \geq \frac{1}{3}$ for all $n \in \mathbb{N}$ since otherwise some player A_j^t could improve her payoff by leaving one of the gadgets $S_{i,\gamma}^t$. Let

$$b_j^n := \Pr_{v_0}^{\bar{\sigma}}(\text{Reach}(F_{B_j^n \bmod 2}) \mid x_n v_n \cdot V^\omega).$$

We have $b_j^n \geq \frac{1}{6}$ for all $n \in \mathbb{N}$ since otherwise some player B_j^t could improve her payoff by leaving one of the gadgets $S_{i,\gamma}^t$. Note that at every terminal vertex of the counter gadgets $C_{j,\gamma}^t$ and $\bar{C}_{j,\gamma}^t$ either player A_j^t or player B_j^t wins. The conditional probability that, given the history $x_n v_n$, we reach one of those gadgets is $\sum_{k \in \mathbb{N}} \frac{1}{2^k} \cdot \frac{1}{4} = \frac{1}{2}$ for all $n \in \mathbb{N}$, so we have $a_j^n = \frac{1}{2} - b_j^n$ for all $n \in \mathbb{N}$. Since $b_j^n \geq \frac{1}{6}$, we arrive at $a_j^n \leq \frac{1}{2} - \frac{1}{6} = \frac{1}{3}$, which proves the claim.

(\Leftarrow) Assume that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. Clearly, this implies that none of the players A_j^t can improve her payoff. To show that none of the players B_j^t can improve her payoff, it suffices to show that $b_j^n \geq \frac{1}{6}$ for all $n \in \mathbb{N}$. But with the same argumentation as above, we have $b_j^n = \frac{1}{2} - a_j^n$ and thus $b_j^n = \frac{1}{6}$ for all $n \in \mathbb{N}$, which proves the claim. Q.E.D.

The second claim relates the probabilities a_j^n and p_j^n .

Claim. Let $j \in \{1, 2\}$. Then $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$ if and only if $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$.

Proof. (\Rightarrow) Assume that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. We have $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ and therefore $\frac{1}{3} = p_j^n + \frac{1}{12}$ for all $n \in \mathbb{N}$. Hence, $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$.

(\Leftarrow) Assume that $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$. Since $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ for all $n \in \mathbb{N}$, the numbers a_j^n must satisfy the following recurrence: $a_j^{n+2} = 4a_j^n - 1$. Since all the numbers a_j^n are probabilities, we have $0 \leq a_j^n \leq 1$ for all $n \in \mathbb{N}$. It is easy to see that the only values for a_j^0 and a_j^1 such that $0 \leq a_j^n \leq 1$ for all $n \in \mathbb{N}$ are $a_j^0 = a_j^1 = \frac{1}{3}$. But this implies that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. Q.E.D.

Finally, the last claim relates the numbers p_j^n to (1).

Claim. Let $j \in \{1, 2\}$. Then $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$ if and only if (1) holds for all $n \in \mathbb{N}$.

Proof. Let $n \in \mathbb{N}$, and let $t = n \bmod 2$. The probability p_j^n can be expressed as the sum of the probability that the play reaches a terminal vertex that is winning for player A_j^t inside C_{j,γ_n}^t and the probability that the play reaches such a vertex inside $C_{j,\gamma_{n+1}}^{\bar{t}}$. The first probability does not depend on γ_n , but the second depends on γ_{n+1} . Let us consider the case that $\gamma_{n+1} = \text{inc}(j)$. In this case, the aforementioned sum is equal to the following sum of two binary numbers:

$$0.00\underbrace{1\dots 1}_{c_j^n \text{ times}}111 + 0.000\underbrace{0\dots 0}_{c_j^{n+1} \text{ times}}100.$$

Obviously, this sum is equal to $\frac{1}{4}$ iff $c_j^{n+1} = 1 + c_j^n$. For any other value of γ_{n+1} , the argumentation is similar, and we omit it here. Q.E.D.

Proof of Lemma 9. By the first claim, the profile $\bar{\sigma}$ is a Nash equilibrium iff $a_j^n = \frac{1}{3}$ for all $j \in \{1, 2\}$ and $n \in \mathbb{N}$. By the second claim, the latter is true if $p_j^n = \frac{1}{4}$ for all $j \in \{1, 2\}$ and $n \in \mathbb{N}$. Finally, by the last claim, this is the case iff (1) holds for all $j \in \{1, 2\}$ and $n \in \mathbb{N}$. Q.E.D.

To establish the reduction, it remains to show that the computation of \mathcal{M} is infinite iff the game (\mathcal{G}, v_0) has a pure Nash equilibrium where player 0 wins almost surely.

(\Rightarrow) Assume that the computation $\rho = \rho(0)\rho(1)\dots$ of \mathcal{M} is infinite. We define a pure strategy σ_0 for player 0 as follows: For a history that ends in one of the instruction gadgets $I_{i,\gamma}^t$ after visiting a black vertex exactly n times, player 0 tries to move to the neighbouring gadget $S_{k,\gamma'}^{\bar{t}}$, such that $\rho(n)$ refers to instruction number k (which is always possible if $\rho(n-1)$ refers to instruction number i ; in any other case, σ_0 might be defined arbitrarily). In particular, if $\rho(n-1)$ refers to instruction $\iota_i = \text{"zero}(j) ? \text{goto } k : \text{dec}(j); \text{goto } l"$, then player 0 will move to the gadget $S_{k,\text{zero}(j)}^{\bar{t}}$ if the value of the counter in configuration $\rho(n-1)$ is 0 and to the gadget $S_{l,\text{dec}(j)}^{\bar{t}}$ otherwise. For a history that ends in one of the gadgets $C_{j,\gamma}^t$ after visiting a black vertex exactly n times and a grey vertex exactly m times, player 0 will move to the grey vertex again iff m is strictly less than the value of the counter j in configuration $\rho(n-1)$. So after entering $C_{j,\gamma}^t$, player 0's strategy is to loop through the grey vertex exactly as many times as given by the value of the counter j in configuration $\rho(n-1)$.

Any other player's pure strategy is "moving down at any time". We claim that the resulting strategy profile $\bar{\sigma}$ is a Nash equilibrium of (\mathcal{G}, v_0) where player 0 wins almost surely.

Since, according to her strategy, player 0 follows the computation of \mathcal{M} , no vertex inside an instruction gadget $I_{i,\gamma}^t$ where ι_i is the halt instruction

is ever reached. Hence, with probability 1 a terminal vertex in one of the counter gadgets is reached. Since player 0 wins at any such vertex, we can conclude that she wins almost surely.

It remains to show that $\bar{\sigma}$ is a Nash equilibrium. By the definition of player 0's strategy σ_0 , we have the following for all $n \in \mathbb{N}$: 1. c_j^n is the value of counter j in configuration $\rho(n)$; 2. c_j^{n+1} is the value of counter j in configuration $\rho(n+1)$; 3. γ_{n+1} is the instruction corresponding to the counter update from configuration $\rho(n)$ to $\rho(n+1)$. Hence, (1) holds, and $\bar{\sigma}$ is a Nash equilibrium by Lemma 9.

(\Leftarrow) Assume that $\bar{\sigma}$ is a pure Nash equilibrium of (\mathcal{G}, v_0) where player 0 wins almost surely. We define an infinite sequence $\rho = \rho(0)\rho(1)\dots$ of *pseudo configurations* (where the counters may take the value ω) of \mathcal{M} as follows. Let $n \in \mathbb{N}$, and assume that v_n lies inside the gadget S_{i,γ_n}^t (where $t = n \bmod 2$); then $\rho(n) := (i, c_1^n, c_2^n)$.

We claim that ρ is, in fact, the (infinite) computation of \mathcal{M} . It suffices to verify the following two properties:

1. $\rho(0) = (1, 0, 0)$;
2. $\rho(n) \vdash \rho(n+1)$ for all $n \in \mathbb{N}$.

Note that we do not have to show explicitly that each $\rho(n)$ is a configuration of \mathcal{M} since this follows easily by induction from 1. and 2. Verifying the first property is easy: v_0 lies inside $S_{1,\text{init}}^0$ (and we are at instruction 1), which is linked to the counter gadgets $C_{1,\text{init}}^0$ and $C_{2,\text{init}}^0$. The edge leading to the grey vertex is missing in these gadgets. Hence, c_1^0 and c_2^0 are both equal to 0.

For the second property, let $\rho(n) = (i, c_1, c_2)$ and $\rho(n+1) = (i', c_1', c_2')$. Hence, v_n lies inside $S_{i,\gamma}^t$ and v_{n+1} inside $S_{i',\gamma'}^t$ for suitable γ, γ' and $t = n \bmod 2$. We only proof the claim for the case that $\iota_i = \text{"zero(2) ? goto } k : \text{dec(2); goto } l\text{"}$; the other cases are straightforward. Note that, by the construction of the gadget $I_{i,\gamma}^t$, it must be the case that either $i' = k$ and $\gamma' = \text{zero(2)}$, or $i' = l$ and $\gamma' = \text{dec(2)}$. By Lemma 9, if $\gamma' = \text{zero(2)}$, then $c_2' = c_2 = 0$ and $c_1' = c_1$, and if $\gamma' = \text{dec(2)}$, then $c_2' = c_2 - 1$ and $c_1' = c_1$. This implies $\rho(n) \vdash \rho(n+1)$: On the one hand, if $c_2 = 0$, then $c_2' \neq c_2 - 1$, which implies $\gamma' \neq \text{dec(2)}$ and thus $\gamma' = \text{zero(2)}$, $i' = k$ and $c_2' = c_2 = 0$. On the other hand, if $c_2 > 0$, then $\gamma' \neq \text{zero(2)}$ and thus $\gamma' = \text{dec(2)}$, $i' = l$ and $c_2' = c_2 - 1$. Q.E.D.

5.2 Finite-state equilibria

It follows from the proof of Theorem 8 that Nash equilibria may require infinite memory (even if we are only interested in whether a player wins with probability 0 or 1). More precisely, we have the following proposition.

Proposition 10. There exists an SSMG that has a pure Nash equilibrium where player 0 wins almost surely but that has no finite-state Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the game (\mathcal{G}, v_0) constructed in the proof of Theorem 8 for the machine \mathcal{M} consisting of the single instruction “inc(1); goto 1”. We modify this game by adding a new initial vertex v_1 which is controlled by a new player, player 1, and from where she can either move to v_0 or to a new terminal vertex where she receives payoff 1 and every other player receives payoff 0. Additionally, player 1 wins at every terminal vertex of the game \mathcal{G} that is winning for player 0. Let us denote the modified game by \mathcal{G}' .

Since the computation of \mathcal{M} is infinite, the game (\mathcal{G}, v_0) has a pure Nash equilibrium where player 0 wins almost surely. This equilibrium induces a pure Nash equilibrium of (\mathcal{G}', v_1) where player 0 wins almost surely.

Now assume that there exists a finite-state Nash equilibrium of (\mathcal{G}', v_1) where player 0 wins with positive probability. Such an equilibrium induces a finite-state Nash equilibrium of (\mathcal{G}, v_0) where player 1, and thus also player 0, wins almost surely since otherwise player 1 would play to v_0 with probability 1. By Lemma 9, this implies that player 0 updates the counter correctly. However, since player 0 uses a finite-state strategy, the corresponding counter values are bounded by a constant, a contradiction. Q.E.D.

Note that FinNE is recursively enumerable: To decide whether an SSMG (\mathcal{G}, v_0) has a finite-state Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$, one can just enumerate all possible finite-state profiles and check for each of them whether the profile is a Nash equilibrium with the desired properties by analysing the finite Markov chain that is generated by this profile (where one identifies states that correspond to the same vertex and memory state). Hence, to show the undecidability of FinNE, we cannot reduce from the non-halting problem but from the halting problem for two-counter machines (which is recursively enumerable itself).

Theorem 11. FinNE is undecidable.

Proof. The construction is similar to the one for proving undecidability of PureNE. Given a two-counter machine \mathcal{M} , we modify the SSMG \mathcal{G} constructed in the proof of Theorem 8 by adding another “counter” (together with four more players for checking whether the counter is updated correctly) that has to be incremented in each step. Moreover, additionally to the terminal vertices in the gadgets $C_{j,\gamma}^t$, we let player 0 win at the terminal vertex in each of the gadgets $I_{i,\gamma}$ where $\iota_i = \text{“halt”}$. Let us denote the new game by \mathcal{G}' . Now, if \mathcal{M} does not halt, any pure Nash equilibrium of (\mathcal{G}', v_0) where player 0 wins almost surely needs infinite memory: to win almost surely, player 0 must follow the computation of \mathcal{M} and increment the new counter at each step. On the other hand, if \mathcal{M} halts, then we can easily construct a finite-state Nash equilibrium of (\mathcal{G}', v_0) where player 0 wins almost surely. Hence, (\mathcal{G}', v_0) has a finite-state Nash equilibrium where player 1 wins almost surely iff the machine \mathcal{M} halts. The details of the construction are left to the reader. Q.E.D.

6 Conclusion

We have analysed the complexity of deciding whether a simple stochastic multiplayer game has a Nash equilibrium whose payoff falls into a certain interval. Our results demonstrate that the presence of both stochastic vertices and more than two players makes the problem much more complicated than when one of these factors is absent. In particular, the problem of deciding the existence of a pure-strategy Nash equilibrium where player 0 wins almost surely is undecidable for simple stochastic multiplayer games, whereas it is contained in $\text{NP} \cap \text{co-NP}$ for two-player, zero-sum simple stochastic games [8] and even in P for non-stochastic infinite multiplayer games with, e.g., Büchi winning conditions [23].

Apart from settling the complexity of NE when arbitrary mixed strategies are considered, future research may, for example, investigate restrictions of NE to games with a small number of players. In particular, we conjecture that the problem is decidable for two-player games, even if these are not zero-sum.

References

1. E. Allender, P. Bürgisser, J. Kjeldgaard-Pedersen & P. B. Miltersen. On the complexity of numerical analysis. In *Proceedings of the 21st Annual IEEE Conference on Computational Complexity, CCC '06*, pp. 331–339. IEEE Computer Society Press, 2006.
2. T. Brázdil, V. Brožek, V. Forejt & A. Kučera. Stochastic games with branching-time winning objectives. In *Proceedings of the 21st IEEE Symposium on Logic in Computer Science, LICS 2006*, pp. 349–358. IEEE Computer Society Press, 2006.
3. J. Canny. Some algebraic and geometric computations in PSPACE. In *Proceedings of the 20th annual ACM Symposium on Theory of Computing, STOC '88*, pp. 460–469. ACM Press, 1988.
4. K. Chatterjee, T. A. Henzinger & M. Jurdziński. Games with secure equilibria. *Theoretical Computer Science*, 365(1–2):67–82, 2006.
5. K. Chatterjee, M. Jurdziński & T. A. Henzinger. Quantitative stochastic parity games. In *Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms, SODA 2004*, pp. 121–130. ACM Press, 2004.
6. K. Chatterjee, R. Majumdar & M. Jurdziński. On Nash equilibria in stochastic games. In *Proceedings of the 13th Annual Conference of the European Association for Computer Science Logic, CSL 2004*, vol. 3210 of LNCS, pp. 26–40. Springer-Verlag, 2004.
7. X. Chen & X. Deng. Settling the complexity of two-player Nash equilibrium. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2006*, pp. 261–272. IEEE Computer Society Press, 2006.
8. A. Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.
9. V. Conitzer & T. Sandholm. Complexity results about Nash equilibria. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI 2003*, pp. 765–771. Morgan Kaufmann, 2003.

10. C. Daskalakis, P. W. Goldberg & C. H. Papadimitriou. The complexity of computing a Nash equilibrium. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, STOC 2006*, pp. 71–78. ACM Press, 2006.
11. L. de Alfaro & T. A. Henzinger. Concurrent omega-regular games. In *Proceedings of the 15th IEEE Symposium on Logic in Computer Science, LICS 2000*, pp. 141–154. IEEE Computer Society Press, 2000.
12. L. de Alfaro, T. A. Henzinger & O. Kupferman. Concurrent reachability games. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science, FOCS '98*, pp. 564–575. IEEE Computer Society Press, 1998.
13. K. Etessami, M. Z. Kwiatkowska, M. Y. Vardi & M. Yannakakis. Multi-objective model checking of Markov decision processes. *Logical Methods in Computer Science*, 4(4), 2008.
14. K. Etessami & M. Yannakakis. On the complexity of Nash equilibria and other fixed points. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2007*, pp. 113–123. IEEE Computer Society Press, 2007.
15. J. Filar & K. Vrieze. *Competitive Markov decision processes*. Springer-Verlag, 1997.
16. M. R. Garey, R. L. Graham & D. S. Johnson. Some NP-complete geometric problems. In *Proceedings of the 8th Annual ACM Symposium on Theory of Computing, STOC '76*, pp. 10–22. ACM Press, 1976.
17. J. F. Nash Jr. Equilibrium points in N -person games. *Proceedings of the National Academy of Sciences of the USA*, 36:48–49, 1950.
18. A. Neyman & S. Sorin, eds. *Stochastic Games and Applications*, vol. 570 of NATO Science Series C. Springer-Verlag, 1999.
19. M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.
20. J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. *Journal of Symbolic Computation*, 13(3):255–352, 1992.
21. R. Selten. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrage-trägheit. *Zeitschrift für die gesamte Staatswissenschaft*, 121:301–324 and 667–689, 1965.
22. M. Ummels. Rational behaviour and strategy construction in infinite multiplayer games. In *Proceedings of the 26th International Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2006*, vol. 4337 of LNCS, pp. 212–223. Springer-Verlag, 2006.
23. M. Ummels. The complexity of Nash equilibria in infinite multiplayer games. In *Proceedings of the 11th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2008*, vol. 4962 of LNCS, pp. 20–34. Springer-Verlag, 2008.
24. W. Zielonka. Perfect-information stochastic parity games. In *Proceedings of the 7th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2004*, vol. 2987 of LNCS, pp. 499–513. Springer-Verlag, 2004.